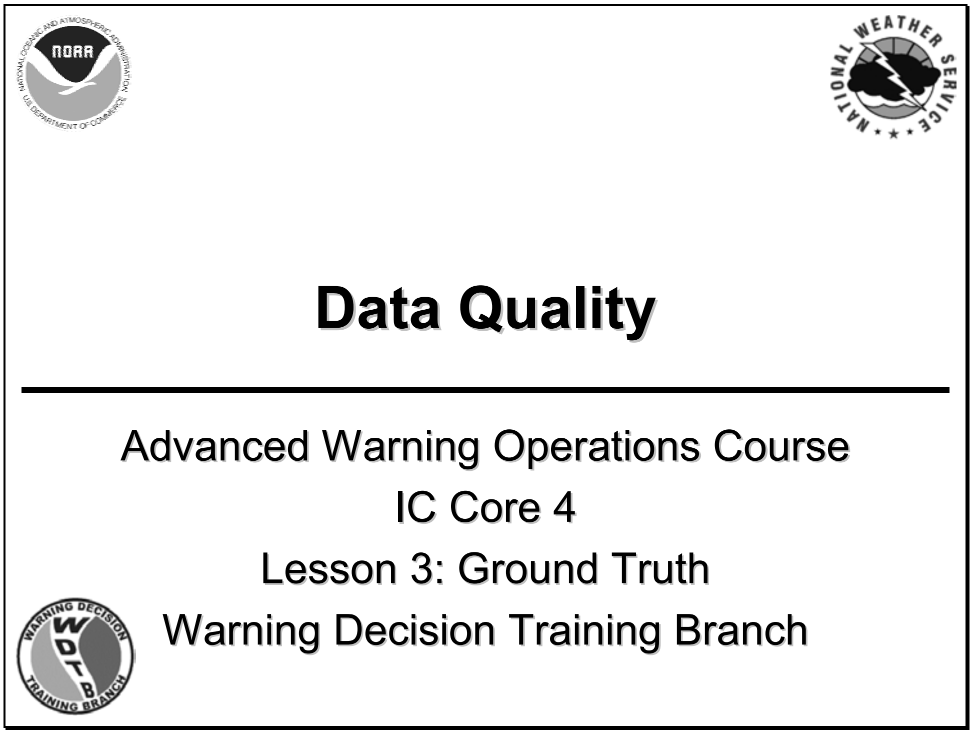
Slide 1



Welcome to the AWOC Data Quality Lesson #3 on Ground Truth data. This presentation should last approximately 25 minutes.

Slide 2

Learning Objectives

1. Identify the sources of ground truth measurements

2. Explain the limitations associated with each type of ground truth measurement

There are two learning objectives for this lesson. First, you should be able to identify the five different sources of ground truth measurements discussed. Next, you should be able list the limitations associated with each of these types of measurement and how they impact the data's effectiveness.

Slide 3

Performance Objective

1. Demonstrate the ability to collect quality controlled ground truth measurements

In addition to the two learning objectives, there is one performance objective for this lesson.

NOTE: Performance Objectives are precise, measurable statements of the behaviors that trainees will be able to demonstrate On-The-Job. They often specify the condition under which the behaviors will be demonstrated as well as the criteria for acceptable performance. (The Performance Objective will NOT be part of the examination process)

The performance objective for this lesson is to demonstrate the ability to collect accurate ground truth measurements during warning operations. Part of the discussion in this lesson will be on some ways to mitigate the impact of poor ground truth measurements through some basic quality control steps.

Slide 4

Topics

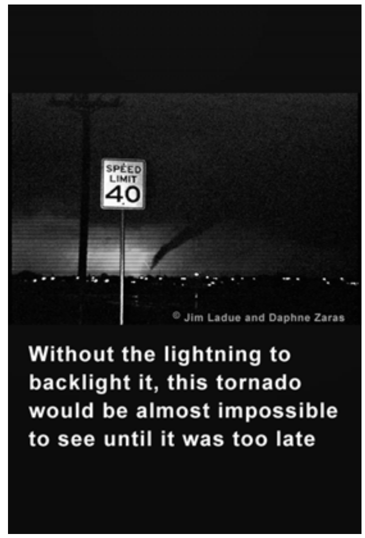
- Spotting storm features and tornadoes
- Hail size
- Wind speed estimation
- Rain gauge measurements
- Snowfall measurements

There are five sources of ground truth data that will be discussed in this lesson. These sources are: spotter-identified storm features (primarily tornadoes), hail size reports, wind speed and damage reports, rainfall measurements, and snowfall measurements. We will next discuss data quality issues and mitigation techniques for each source of data.

Slide 5

Spotting Storm Features and Tornadoes

- Cloud features misidentified by some
- Problem magnified at night
- Lightning illumination
- Power flashes

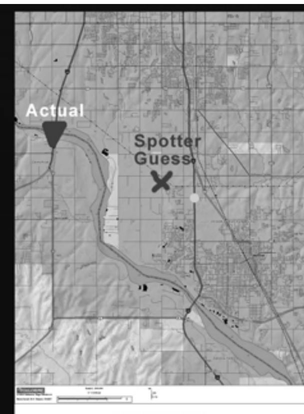


Inexperienced storm spotters can often misidentify cloud features. These well-meaning folks may interpret low-hanging scud cloud as a rapidly rotating wall cloud, funnel, or even a tornado. While educating people helps with this problem during the day, its a bigger problem at night. In low-lighting conditions, even expert spotters may have trouble identifying storm features. Many times the best source of light will be from lightning. Another potential identifier of a nocturnal tornado are power flashes. Just remember that power flashes may also occur with strong straight line winds, even strong inflow winds.

Slide 6

Spotting Storm Features and Tornadoes (Cont)

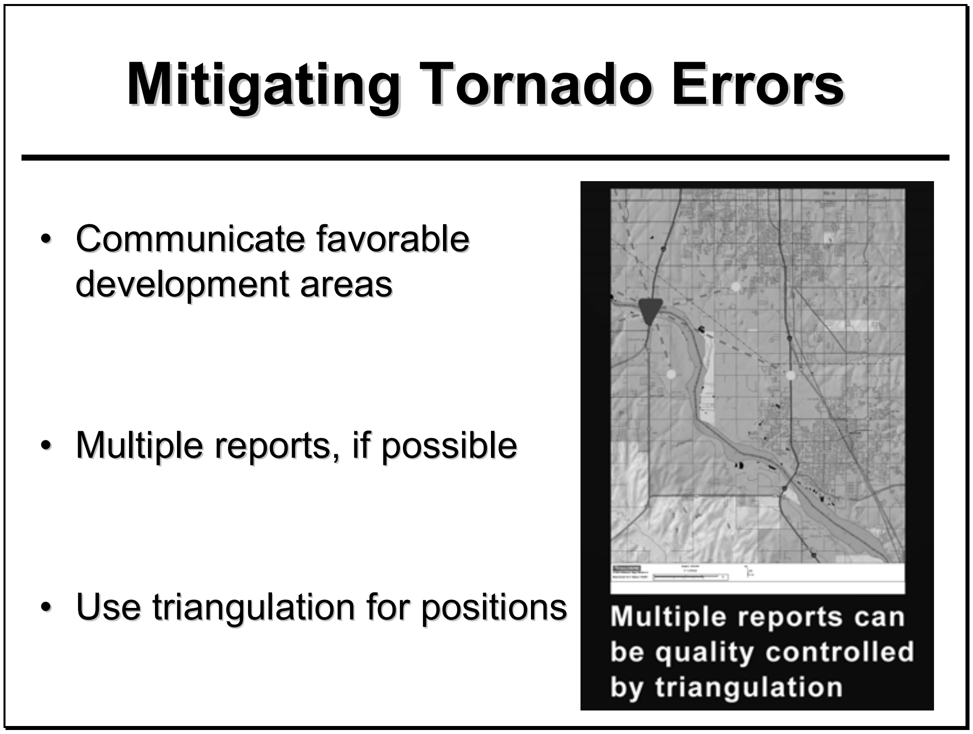
- Difficult to gauge distance
- Lack of reference points
- Actual distance double (or more) of estimate



A single spotter report can be hard to quality control

Besides spotting a storm feature, estimating distance to a feature can also be a problem. Objects in the sky at a distance tend to appear closer than they actually are. The problem is due to a lack of intermediate reference points. For example, say a spotter only provides an estimated tornado location. If that estimate is off significantly, it can cause problems. However, if you also know the spotter’s location, or even which direction they are looking, then that additional info can help if the estimate is incorrect. After all, the actual distance to the object that the spotter is seeing may be two or more times the estimated distance. No one wants to send the wrong group of people into their storm shelters.

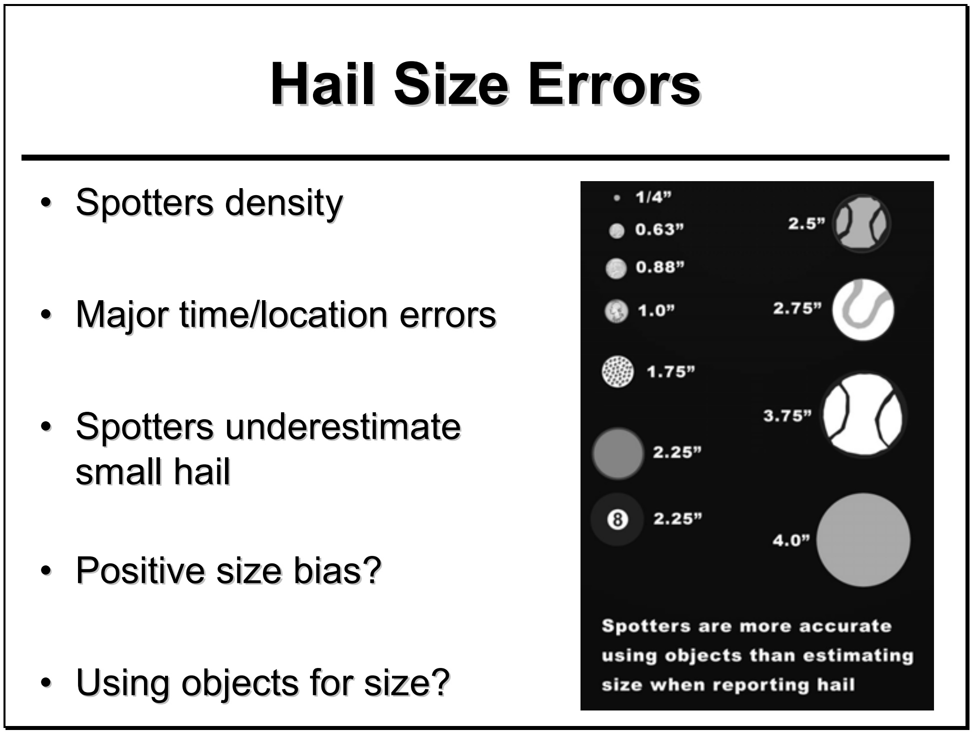
Slide 7



To mitigate these errors and their impact, one thing you can do is to effectively communicate areas that are favorable for tornadic development to spotters. You don't want to come right out and say you are looking for information about a possible tornado in such and such location. What you do want is to communicate to spotters where the storms are and where the biggest threat may occur without leading spotters in a way that results in a self-fulfilling forecast. Doing so will help spotters focus on key areas and should help eliminate spurious reports.

Another way to help eliminate some errors is to get as many reports of a specific tornado as possible. If you do get a questionable location, having multiple reports of the tornado, funnel cloud, or wall cloud will help you triangulate (Speheger, 2004), as in the example shown. Even comparing reports of a tornado with radar data can help clarify the tornado's location.

Slide 8



There are a variety of issues with hail size. While the studies may seem contradictory at times, the overall conclusions that can be drawn are consistent.

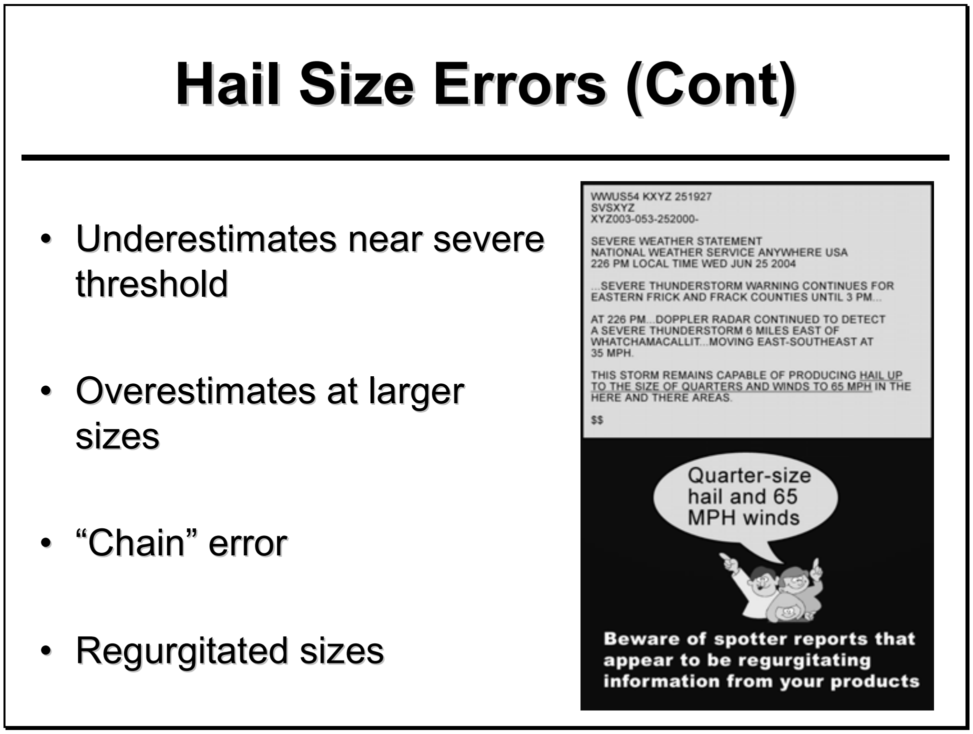
Changnon (1968) determined that you need an observing network with a density of one observer per square mile to accurately measure hail size. Most spotter networks are less dense than that, which introduces error.

Witt et al. (1998) found that a small sample (115) of hail reports from OK and FL had an error rate of about 30% for time or location. This percentage was based on hail reports being located on the edge of a cell or over 50 km away from the radar-defined storm. (The same study used a range of 30 km to associate a hail report with a storm) While a detailed study has not been conducted, in talking with various experts it's believed that a more representative percentage for a large sample size of storm reports (wind and tornado included) would be closer to 10%.

Herzog and Morrison (1994) state that there is evidence that hail sizes in Storm Data have a substantial bias towards larger hail stones. However, Baumgardt et al. (2001) found that, lacking an objective measuring tool, spotters tend to underestimate hail smaller than golf ball size. At golf ball size and larger, the bias is less significant, but the standard deviation in estimated sizes increases. This result suggests that, as hail gets larger, it is easier to get an overestimate in hail size.

The same group found in a later study that spotters are more accurate identifying hail size by comparable object (egg, golf ball, etc.) than by objective measure (one inch, half-inch, etc.). While this result would seem to help, Edwards and Thompson (1998) shows a graphic that suggests that hail reports tend to accumulate at sizes corresponding to certain objects, including golf ball and baseball sized hail. In addition, the climatological record contains some errors with regard to hail being compared to certain sized objects that do not match the measured amount listed (i.e., softball size hail is often listed as 4.5”, but the largest regulation softball is 3.75”).

Slide 9



Based on these previous research studies here's how hail sizes may impact operations.

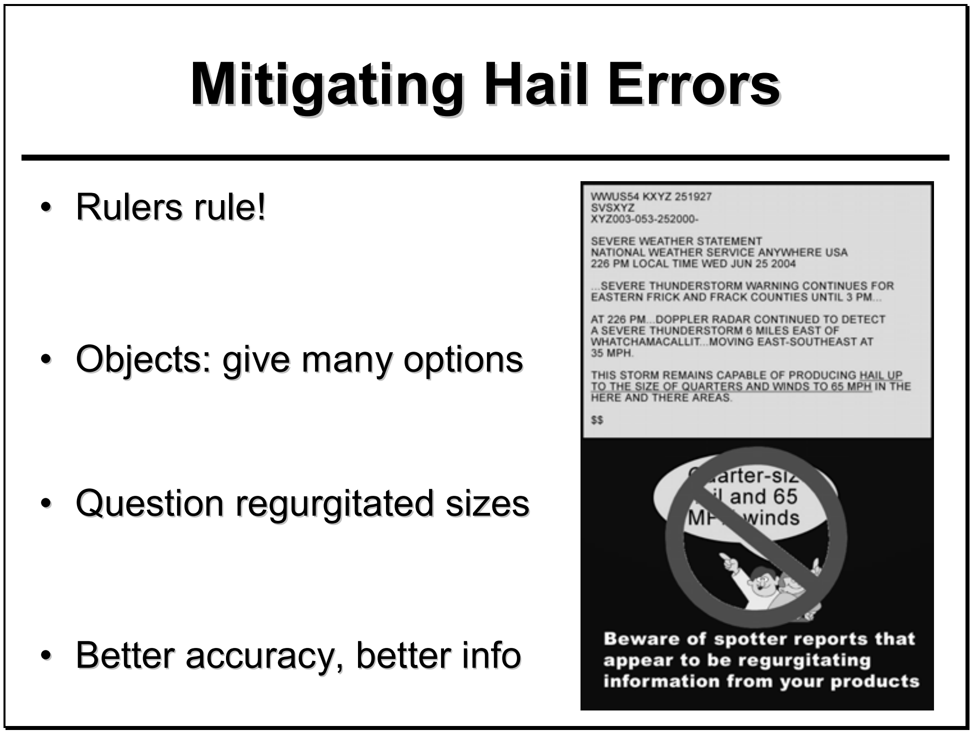
The general public and hasty spotters will tend to underestimate the size of small hail. When I say small, I mean smaller than golf ball size. However, the significance of the underestimation is really at 1" or smaller since the severe criteria is ¾".

Around golf ball size and larger, overestimation is a more significant problem. At golf ball size and smaller, ~75% of estimates are smaller than the actual size. By 2.5", it's down to around 55-60% of estimates are below the actual size (Baumgardt et al., 2001). These distributions were based on numerical measurement estimates. Another impact at this range of sizes is with comparable objects. Below golf ball size, there are numerous coins and other objects commonly used to identify hail size. At golf ball size and larger, there are fewer objects that people can identify with a hail stone. This would explain the clustering of hail size distributions around golf ball, baseball, and softball size hail that was seen in Edwards and Thompson (1998).

Many times the spatial and temporal errors with hail reports are due to "chain" error. The more people that handle a particular spotter report, the greater likelihood there is going to be an error associated with that measurement. Some of the size (as well as time and location) errors seen may be attributed to this problem.

Yet another possible source of size bias is when a report repeats the hail size mentioned in NWS products, media broadcasts, or algorithms. While we want to provide the general public and our other customers some kind of an estimate of the threat from potential hail, it's important to pay attention to how closely the reports coming in match the forecast hail size. If you are getting multiple reports of the same size hail, and just happens to match your products, you might want to view the reports cautiously. There likely is hail at those locations, but the size estimates may be way off. This phenomenon may balance out some of the biases discussed previously and help make the research findings the murky picture that they are.

Slide 10



Experienced spotters are very good about carrying rulers with them to measure hail. Those reports will be the most accurate.

If a spotter doesn't have a ruler, try to get them to compare the hail size to an object. In doing so, try to give them several options. This process is especially important at larger hail sizes. There's a big gap between a golf ball and a tennis ball, or a baseball and a softball!

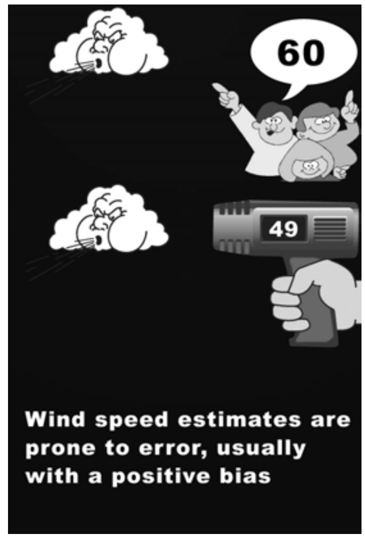
If you receive hail sizes that appear to repeatedly match forecast sizes, algorithm sizes, etc., rely on sizes provided by trusted spotters. You should expect some variability in your reports that come in, but a measurement from a proven spotter should dismiss concerns of inaccuracy.

While some of these steps may be time consuming, their goal is to provide you with more accurate, and better, information.

Slide 11

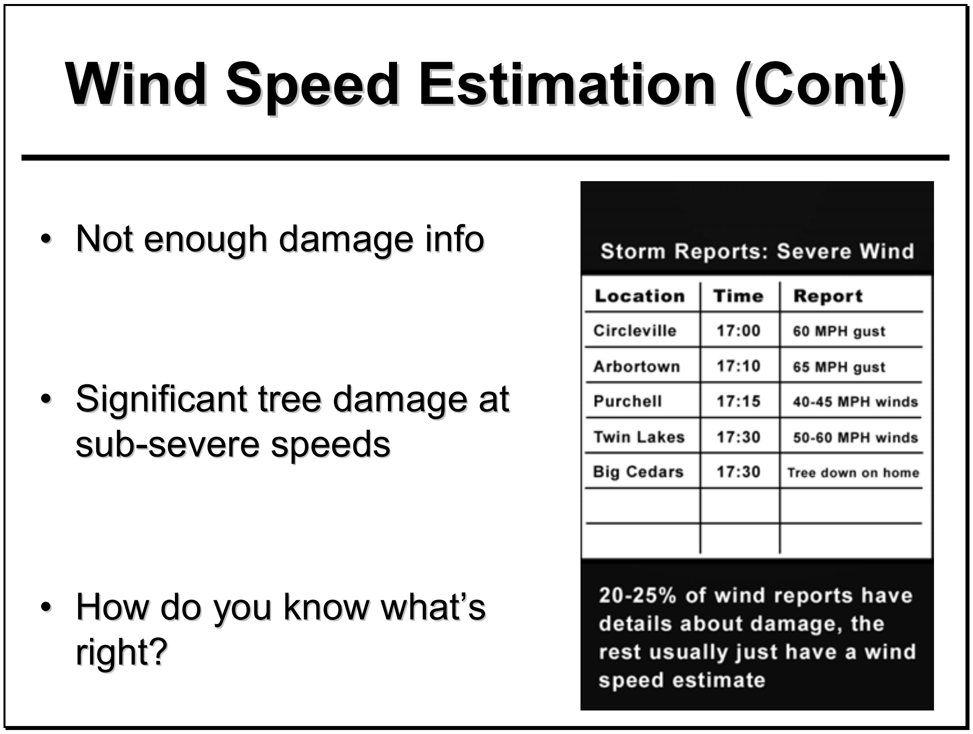
Wind Speed Estimation

- Most prone to error?
- Overestimated, generally
- More marginal events?



Of all of the information that forecast offices receive from spotters, wind speed estimates may be the most prone to error. Some reports are based on damage while others are based on spotter-derived estimates. These estimates are very subjective! While spotters do their best, they tend to overestimate wind speed (LaDue, 2003). This problem can be a particular problem at marginally severe or sub-severe speeds since more attention has been paid to observing these events over the last couple of decades (Weiss et al., 2002).

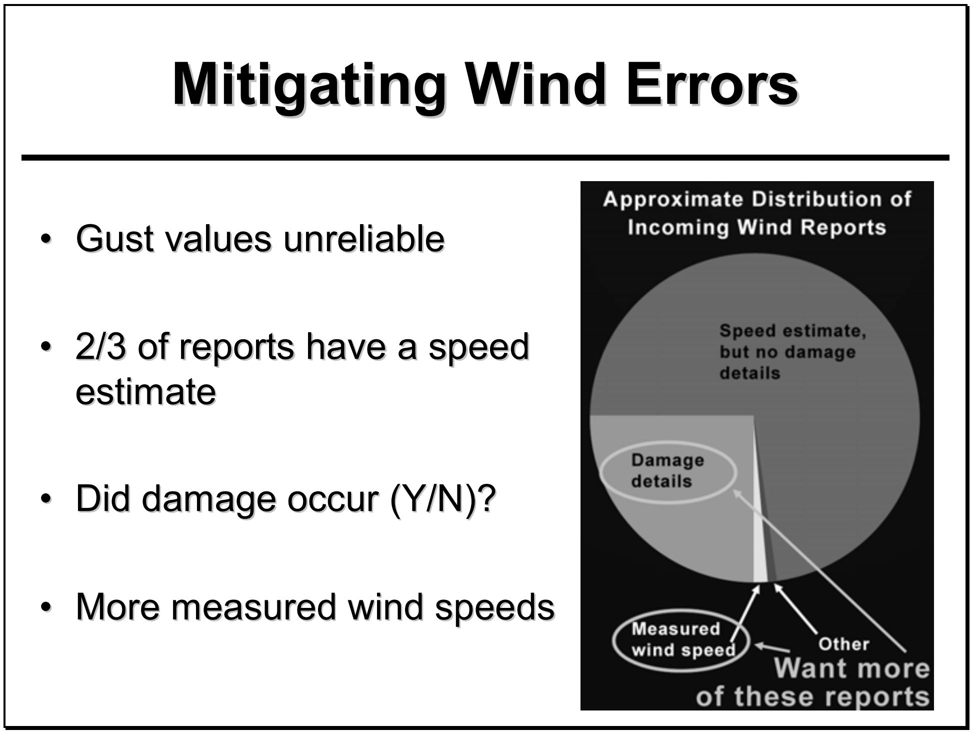
Slide 12



In the cases where damage is reported, there is at least some objective result to the wind event. The problem is that wind damage is often not well reported (or not well documented before it gets to the warning forecaster). One study (Weiss and Vescio, 1996), found that ¾ of all thunderstorm wind reports over a 30-year period were listed as "wind damage" with little or no detailed information about the type of damage. In a review of a small sample of events, we found that reports in phone logs generally had some flavor of wind damage listed in only 1 in 5 reports.

Even if damage is reported, much of the wind damage reports are characterized by tree damage. It's doubtful that many spotters know much all the factors that may impact how strong winds have to be to do tree damage. In some areas of the country, significant tree damage can occur at wind speed below the severe critieria. The point is that it's very easy for spotters to misjudge the intensity of thunderstorm winds. How are warning forecasters supposed to know if spotters are getting it right?

Slide 13



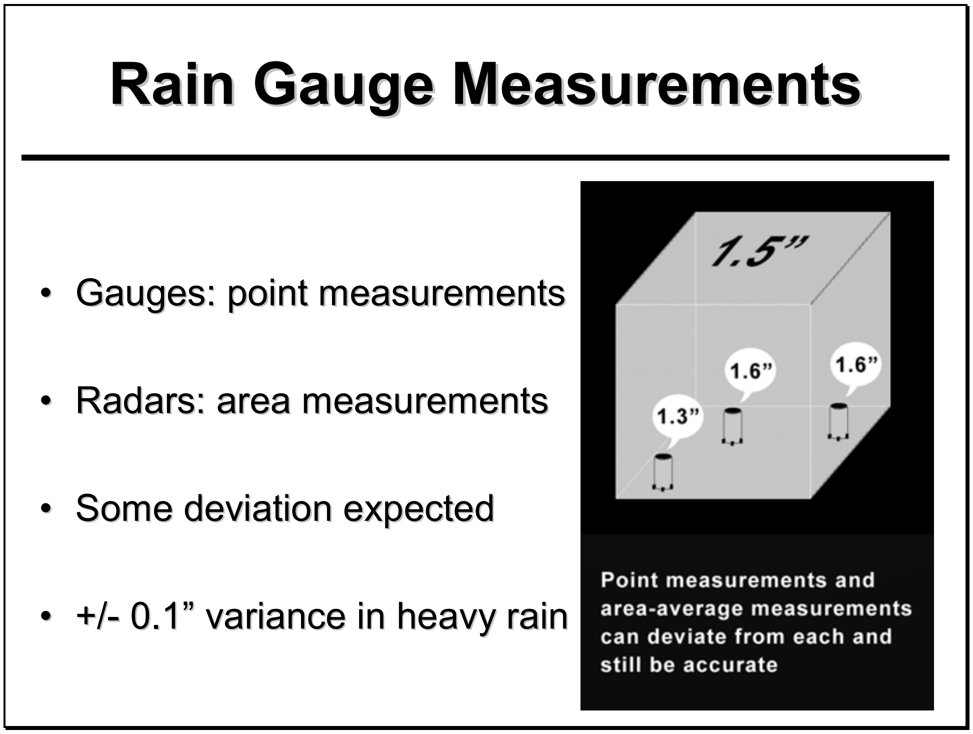
While wind gust estimates may not always be accurate, they do have value. Much like hail reports, forecasters should be aware of people seeming to repeat back forecast wind speed values from your products.

In a review of some previous events, it was found that about 2/3 of wind reports included some kind of wind estimate. Instead of focusing on the numerical value of the wind estimate, it might be better to look at the range of estimates relative to the severe threshold. Say something like 0-20 (light), 20-45 (strong, but definitely sub-severe), 45-65 (marginally sub-severe to marginally severe), 65-80 (definitely severe), and 80+ (take cover now). The numbers will, and should, vary depending on your CWA. One key to such a system is to have a good idea at what wind speed tree and other wind damage occurs in your CWA. Why you ask? So you can use reports of wind damage to QC the wind estimates.

Of course, you can't use wind damage to QC estimates if you don't receive reports of damage. When a report comes in, especially if you are receiving a first hand report, always ask if there was any damage. Make sure you document the answer! You are likely to forget in a few minutes if you have a lot of reports coming in.

Another good way to QC wind estimates, if possible, is to use measured wind speeds from METARs, mesonet sites, or spotters (via portable or hand-held anemometers) located near by. It is helpful if peak gust values from these sites can be recorded on spotter log sheets as storms move through. Even if the data comes in minutes (even an hour) later, it helps to get that information in your logs. You will want that information for any post-mortem event review that you perform.

Slide 14



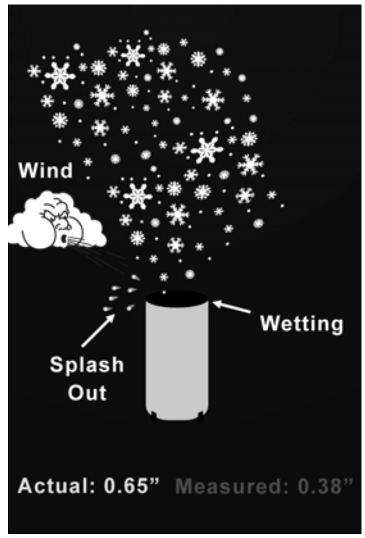
When errors associated with rainfall totals are usually mentioned, the focus is on radar measurement errors. Here will take a few moments to discuss rain gauge measurement errors and the acceptable variations between point and area-averaged data. While the errors with rain gauges are usually smaller than that with the radar estimates, they can still be significant. An important point to remember when talking about rainfall totals is the difference between the two data sources. Rain gauge rain totals are point measurements. They tell you the rainfall amount at a given point in space. The radar rain total is an area-averaged measurement. These data tell you the rainfall amount over an area defined by the radar specifications and algorithms. Some deviation between the two values is to be expected, especially under certain circumstances.

The previous biases really just impact the difference between the point measurement at a rain gauge and what the rain gauge should have measured. There is also the difference between what a point gauge measures and the area-averaged value for a radar measurement. In heavy rain, the standard deviation for a point measurement vs. the area average can be 0.05” or more (Duchon et al., 1995). That means that it is not unusual for accurate rain gauge measurements to deviate from a radar measurement by up to 0.1” or more just based upon the difference type of measurement.

Slide 15

Rain Gauge Measurements (Cont)

- Undercatch: wind, splash out, and wetting
- Tipping bucket gauge issues
- Rain: 3-10% low bias
- Snow: 50% or higher!

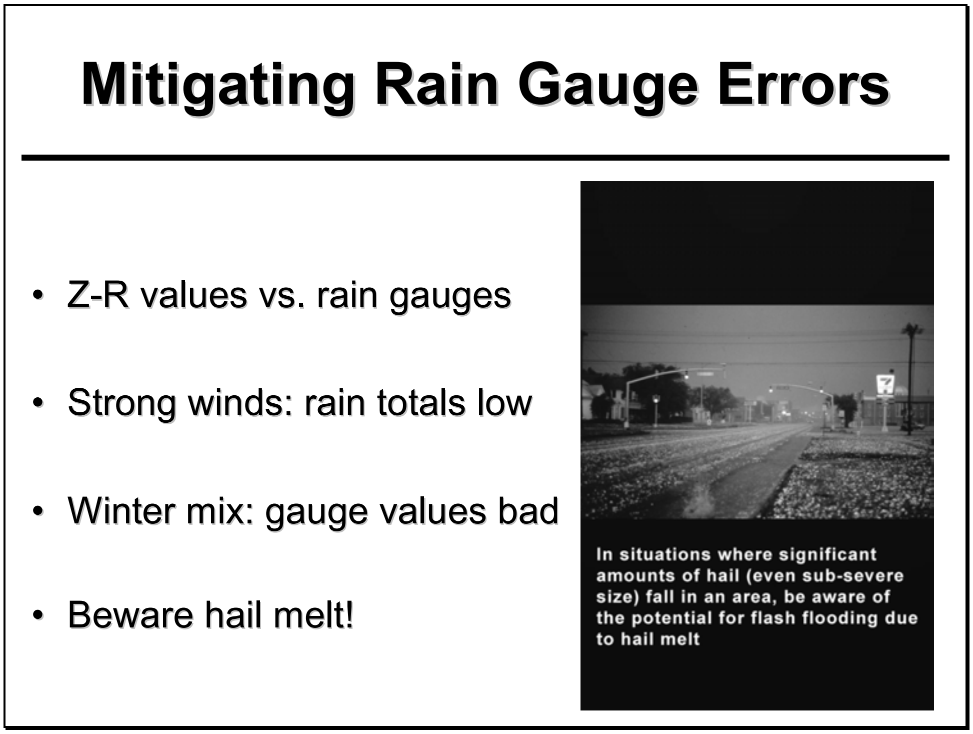


An important thing to remember about rain gauges is that they primarily have a negative bias. The reason for this tendency is that most of the error sources in a rain gauge result in less rain being measured by the gauge than what actually has fallen. These errors sources include wind-induced undercatch, heavy rain splash out, and gauge wetting. Undercatch is an issue that effects just about ever rain gauge, including ones with features (like wind and splash guards) used to minimize them.

Another source of undercatch not in the above list occurs with tipping bucket gauges. In heavy rain, tipping bucket gauges measure less rainfall than a standard rain gauge by up to 12% (Trammel, 2004). The undercatch is a result of some precipitation being lost in the process of the buckets tipping. Another issue with tipping bucket gauges, double tipping, can actual result in a positive bias. Double tipping can occur if a rain gauge is unevenly sited, which is extremely rare for most NWS observing sites. However, it is a possibility with some COOP or spotter sites, if they are using an automated, tipping bucket rain gauge (which is uncommon). Because the positive bias will generally be very rare, the discussion will focus on the low bias issues.

In general, the low bias for rain only events is approximately 3-10% (Groisman and Legates, 1994). While this value may not seem like much, it can be significant when you are talking about heavy rainfall occurring in a short time span. For snow and wintry mixes, the low bias is significantly larger. It can be 50% or higher! This bias makes most rain gauge measurements of liquid equivalent useless during a weather event that has winter precipitation for any significant amount of time.

Slide 16



Here are a few things to remember when dealing with possible rain gauge errors:

Don't be surprised if Z-R rainfall values end up being a little high, especially in heavy rainfall areas. In areas that have received 2-5" of rain, rain gauge measurement errors could account for an undercatch of .25-.5". Even if radar estimated rainfall rates are poor, warning forecasters should realize that the rain gauge measurement may have errors, also. If measured rainfall amounts approach 90% of flash flood guidance, there is a possibility that the actual rainfall amounts could be approaching 100%.

When storms with strong winds (especially severe thunderstorm intensity winds) move through an area, rain gauge measurements should be expected to be low. While it's difficult to say the exact bias for each site without detailed study of the instrument, 10% may be a good guess. Besides the fact that it falls in the range of rain gauge bias values, it's probably one of the easier values to figure out in a time sensitive situation. You just want to come up with a number that you feel comfortable with as being accurate enough to make a good warning decision.

In winter warning situations, it's best to avoid rain gauge totals from any site that has received any significant snowfall during the measuring period. When trying to determine liquid equivalent rain measurements, you will almost always get more accurate values from snowpack measurements than from rain gauge measurements.

One last issue about mitigating rain gauge errors is with regards to hail melt. If you have a significant hail storm in your county warning area that results in hail covering the ground, forecasters should beware of the impact hail melt will have on flood prone areas. This problem is especially true in urban areas. Since little, if any, hail will accumulate in a rain gauge, the measured rainfall total is likely to be a lot less than the actual rain and hail liquid equivalent that has fallen in that area. Combine that deficiency with the fact that melting hail has a habit of clogging man-made drainage areas and you can have a major problem in low-lying areas.

Slide 17

Snowfall Measurements

- Climatological vs. real-time
- Timing of observations key
- Climate: every 6 hrs to 1/day
- Real-time: every 1-6 hrs

Real-Time Observations

Location	Time	Amount
Circleville	7:00 AM	1”
Circleville	8:00 AM	0.5”
Circleville	9:00 AM	2.0”
Circleville	11:00 AM	3.3”

Climatological Observations

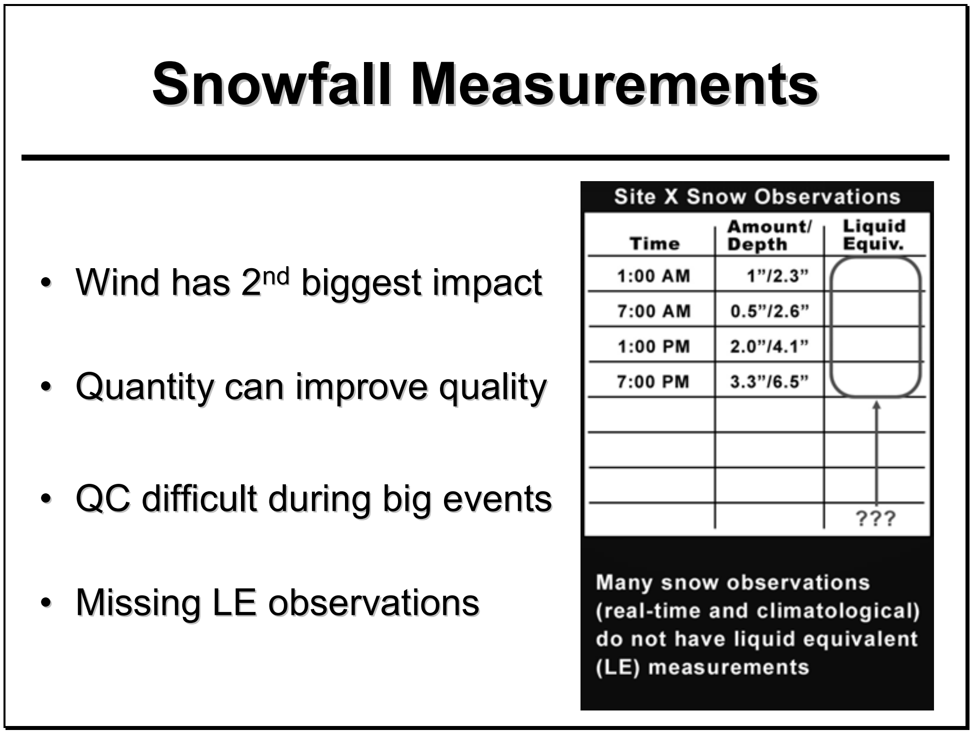
Location	Time	Amount
Circleville	6:00 AM	5.1”
Circleville	12:00 PM	1.1”
Circleville	6:00 PM	0.3”
Circleville	12:00 AM	0.0”

When discussing snowfall measurements for ground truth purposes, it’s important to discuss the two different types of snowfall measurements available to forecasters: climatological and real-time measurements. The difference between the two measurements is primarily in the timing of the observations (NOAA, 1997).

Climatological snowfall observations, which most of you are familiar with, are taken anywhere from once a day to every 6 hours. The reports include (or should include) new snowfall, snow depth, and liquid equivalent. These reports are used as part of the climatological record of snowfall.

Real-time snowfall observations occur more frequently than climatological observations, usually on the order of every 1-6 hours. The observations are useful to the warning forecaster to help them determine intensity and duration of an event in a particular area and issue products accordingly. However, because these reports are taken more frequently than every 6 hours, the snow doesn’t have enough time to settle so these measurements are not accurate enough for snow total measurements.

Slide 18



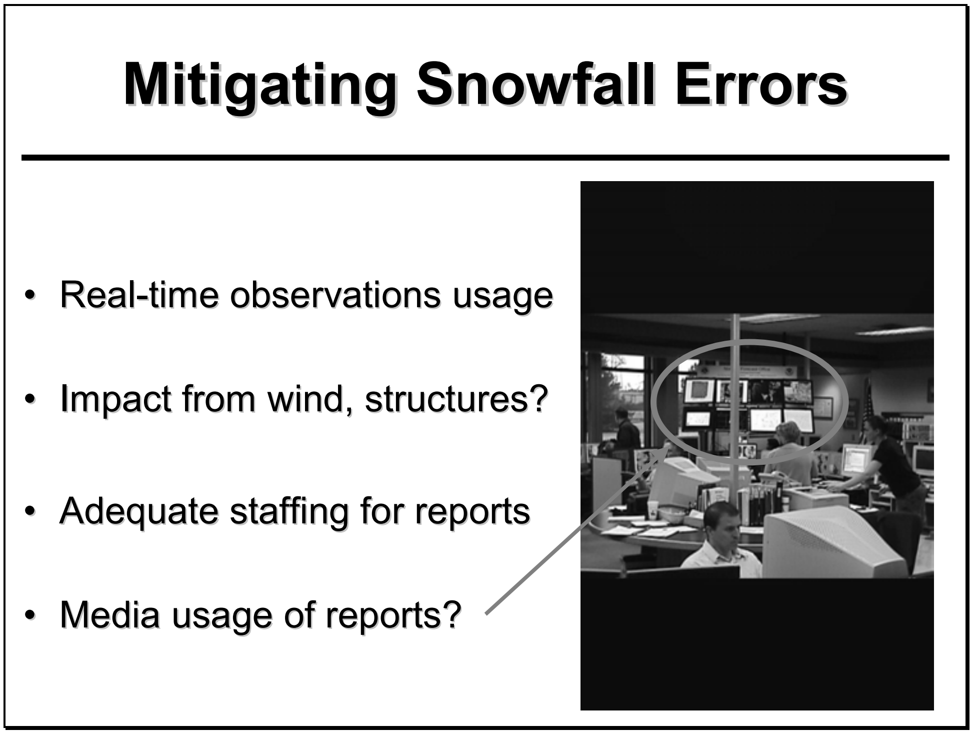
Besides the frequency of observations, one of the biggest impacts on snowfall measurement accuracy is wind. While sites are usually picked to minimize the impact of wind, that is not always an option. Also, strong winds (i.e., blizzard conditions) can still have an impact on drifting at sites that are selecting specifically to minimize the effect of drifting. In some regions of the country, where the land is generally flat and lacks sufficient tree coverage, it is virtually impossible to get accurate snowfall totals during blizzard conditions.

It is important, as a result, to have a good quantity of observers available to you. The volume of information may at times be difficult to handle. When you are impacted by a mesoscale winter weather event, having accurate, detailed real-time observations of the event may be the only good way to keep on top of the event and maintain good SA.

The volume of these reports, if you have a large number of spotters, may make QC of incoming reports difficult. This problem is worse the more real-time observations you receive because of the shear volume of reports (BUF office had about 15 real-time observers per county, on average, back in 1997; that number may be closer to 20 by now).

While real-time reports may focus primarily on intensity, liquid equivalence (LE) is important, too. LE is a variable that is not impacted by the frequency of observations. However, climatological reports (1-4 a day) often do not contain LE amounts, let alone reports that are coming in every hour or so. Many times, LE reports from spotters may be the only reliable means for forecasters to know the amount of precipitation that has fallen (due to unreliability of rain gauges during winter weather).

Slide 19



Based upon my previous statement, it's clear that have real-time snow spotters is a big help. In a service assessment of a significant lake effect snow storm in western NY, one of the recommendations was for forecast offices to develop a network of real-time snow spotters to work in conjunctions with climatological snowfall observers. The real-time observers have had a major impact in the BUF CWA during several events since then (NOAA, 1997).

It's important to note again that real-time observations of snowfall are good at determining snowfall intensity, not snowfall totals. Most times snowfall totals from hourly measurements will exceed climatological measurements because the snow has not had a chance to settle and compact during the hourly measurements.

Just like with severe storm spotters, communication is always important. The more information you can get about the observation the better. Were the winds strong at the measurement site? Are there any structures nearby that may have had an impact on the measurement? Anything you can learn (and document) that might have a data quality impact on the measurement is important.

Of course, to handle this flow of information you will need to make sure there is adequate staffing for taking and quality controlling the reports.

You will also need staffing to handle one potential side effect of gathering these real-time snowfall observations: media reports. It's possible the media, not knowing the difference between climatological and real-time snowfall measurements, will sum up the real-time snowfall measurements and report the totals to the public. If there is a significant deviation between the official totals you report and the media's report, you may get phone calls from confused people wondering which number is correct.

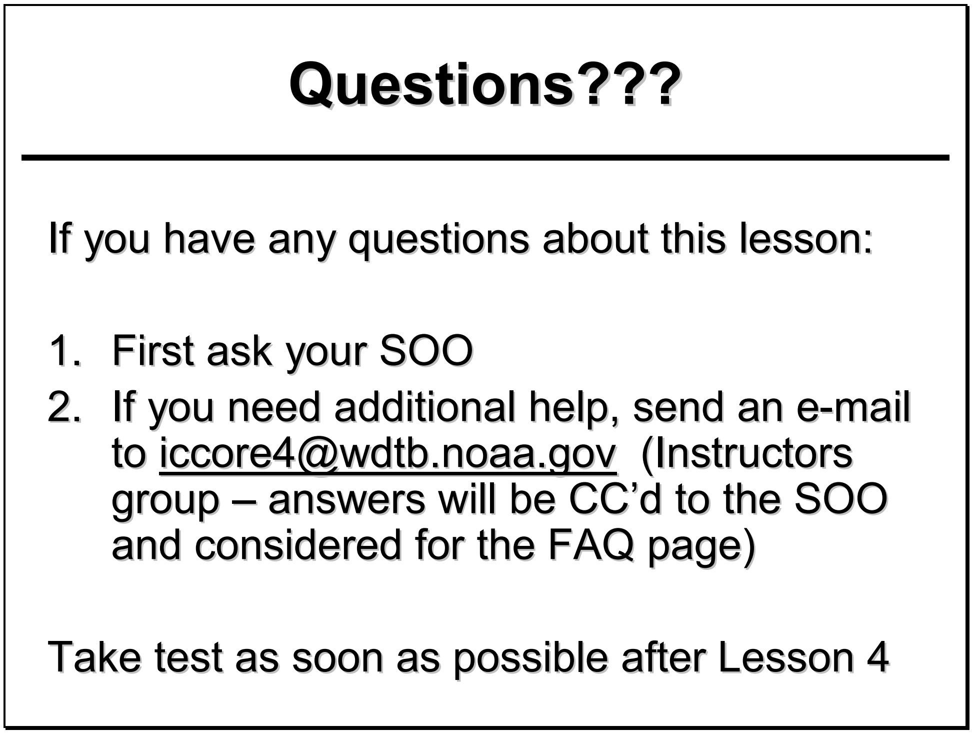
Slide 20

Conclusions

- Understand ground truth data quality issues and mitigate when possible
- You can find some, but not all errors, during warning operations
- Goal is to find the obvious errors, corroborate questionable reports
- Adequate staffing, good communication key

Regardless of the type of ground truth data, the most important thing is to understand the issues that can affect data quality and know how to mitigate those issues when possible. Many times, you will not be able to prevent the error from occurring, but you can prevent the bad data from impacting your operations. While some ways to mitigate data errors are provided here, it’s important to realize that no technique will be perfect. You will not be able to catch every bad ground truth report. With a little hard work, however, you should be able to determine if a report has an obvious error. With a little luck, you may even be able to corroborate or dismiss reports that you receive and identify as questionable. All you are trying to do is get the best information possible. To take these steps, it’s important to make sure you have an adequate number of staff available to you during warning operations and that good communication is maintained prior to and during the event with the people who are making the ground truth observations for you.

Slide 21



If, after going through this lesson you have any questions, first ask your SOO. Your SOO is your local facilitator and should be able to help answer many questions. If you need additional info from what your SOO provided, send an e-mail to the address on the slide. This address sends the message to all the instructors involved with this IC. Our answer will be CC'd to your SOO so that they can answer any similar questions that come up in the future. We may also consider the question and answer for our FAQ page. Thanks for your time and good luck on the exam!

References

Speheger, D. A., 2004: Personal communications.

Davies, J. M., C. A. Doswell III, D. W. Burgess, and J. F. Weaver, 1994: Some noteworthy aspects of the Hesston, Kansas, tornado family of 13 March 1990. *Bull. Amer. Meteor. Soc.*, **75**, 1007-1017.

Chagnon, S. A., 1968: Effect of sampling density on areal extent of damaging hail. *J. Appl. Meteor.*, **7**, 518-521.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286-303.

Witt, A., M. D. Eilts, G. J. Stumpf, E. D. Mitchell, J. T. Johnson, and K. W. Thomas, 1998: Evaluating the performance of WSR-88D severe storm detection algorithms. *Wea. Forecasting*, **13**, 513-518.

Herzog, R. F., and S. J. Morrison, 1994: Hail frequency in the United States, Haag Engineering Co. Rep., 18 pp. [Available from Haag Engineering Company, P.O. Box 814245, Dallas, TX 75381.]

Edwards, R. and R. L. Thompson, 1998: Nationwide comparisons of hail size with WSR-88D vertically integrated liquid water and derived thermodynamic sounding data. *Wea. Forecasting*, **13**, 277-285.

Baumgardt, D., D. Rosendahl, and T. Shea, 2001: The Storm Spotter Hail Approximation Project: How good are your spotters?. NWS, LaCrosse, http://www.crh.noaa.gov/arx/.

LaDue, J. G., 2003: The role of spotters in the integrated warning system. Sierra Storm Broadcasters Conference, Reno.

Weiss, S. J., J. A. Hart, and P. R. Janish, 2002: An examination of severe thunderstorm wind report climatology: 1970-1999. Preprints, 21st Conf. on Severe Local Storms, San Antonio , Amer. Meteor. Soc., 446-449.

Weiss, S. J., and M. D. Vescio, 1996: Severe local storm climatology 1955-1996: Analysis of reporting trends and implications for NWS operations. Preprints, 19th Conf. on Severe Local Storms, Minneapolis, Amer. Meteor. Soc., 536-539.

Duchon, C. E., T. M. Renkevens, and W. L. Crosson, 1995: Estimation of daily area-average rainfall during the CaPE experiment in Central Florida. *J. Appl. Meteor.*, **34**, 2704-2714.

Trammel, 2004: Personal communications.

Groisman, P. Ya., and D. R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteor. Soc.*, **75**, 215-227.

NOAA, 1997: Evaluation of the reported January 11-12, 1997, Montague, New York, 77-inch, 24-hour lake-effect snowfall. pp. 46.